Stops Report

Executive Summary

This report analyzes 2024 Cleveland Department of Police Stops data by race and gender. It uses a pipeline lens, tracking encounters from the initial stop through enforcement outcomes. Findings are shown separately for Traffic Stops and Investigatory Stops.The analysis distinguishes between raw disparities and disparities that remain after accounting for observable context (e.g., district/zone, stop reason, time, age, subject role, stop type). It also includes bias test analysis - such as hit-rate and threshold tests - to help separate contextual variation from potential systemic differences in process.

After accounting for observable situational factors, the report finds significant Black - White differences in search rates across both Traffic and Investigatory Stops. In Traffic Stops, Black subjects are searched 50.5% more frequently than White subjects after controlling for factors such as district and zone of encounter, time and year of encounter, reason for the stop, subject role, and age. In Investigatory Stops, Black subjects are searched 13.5% more often than White subjects in comparable contexts. Even so, the report does not find evidence of process bias in search decisions using the contraband hit-rate and threshold tests for these core comparisons, and post-search outcomes are broadly comparable by race in the headline results presented, including similar contraband hit-rates and arrests conditional on search. At the district level, the contraband hit-rate test likewise does not flag significant differences in any district once Bonferroni-adjusted p-values are used to correct for multiple hypothesis testing.

Female subjects are searched substantially less than Male subjects in both Traffic Stops (controlled −47.1%) and Investigatory Stops (controlled −30.0%), but search decisions do not fail the bias tests in those comparisons. One notable exception for post-search outcomes appears after a Traffic Stop search: conditional on being searched, Female subjects are 29.3% less likely to be arrested, particularly in District 5.

This report analyzes the department’s administrative stop-and-search records as recorded. As a result, it can describe patterns in stops, searches, and arrests, but it cannot fully explain the decision-making behind them. Two limitations are central: (1) the data cannot reliably distinguish discretionary searches from non-discretionary searches required by policy or procedure, which is where bias is most directly evaluated, and (2) the dataset lacks structured detail on search justification (e.g., probable cause or reasonable suspicion) and in-stop dynamics (e.g., behavior, observations, escalation), so unobserved factors may still drive part of the differences we observe. Finally, because the data do not include downstream indicators of arrest quality (e.g., case disposition or prosecution outcomes), arrest comparisons should be interpreted as differences in observed arrest rates – not as direct measures of arrest justification or efficiency.

It is of paramount importance to understand whether policing is applied fairly, not just effectively, because legitimacy and public trust depend on it. This report attempts to do that by applying statistical rigor – moving beyond topline averages to test whether disparities persist in comparable contexts, reflect legitimate enforcement objectives versus demographic proxies, and provide an evidence-based foundation for procedurally just interventions.

Table of Contents

The Data

This report analyzes Cleveland Department of Police Stops data across 2024, transformed into a single analysis-ready table with standardized stop, search, contraband, outcome, location, time, and demographic fields.

Key analytic measures include whether a stop involved a search (subject and/or vehicle), whether any contraband was seized, and whether the stop resulted in an arrest or other dispositions like citations and warnings.

Demographics are standardized (e.g., gender/race/ethnicity mappings) and age is validated to reduce data-entry errors.

The raw data contain some duplication and validity issues that are related to the way data are stored and classified. For example, incident IDs and citation IDs are not always unique, and data important to establishing stop justification or context contain non-standardized, free text fields. These issues require steps to reduce duplication and validity issues in the final dataset that are important to be aware of. More details on the transformation and deduplication process can be found in Appendix A.

A Note on the Data's Limitations

This report uses administrative stop-and-search records to describe patterns by race and gender. These data are useful for measuring what happened (e.g., who was searched, whether contraband was found, whether an arrest occurred), but they do not capture every detail needed to explain why a search or arrest happened. The limitations below are important for interpreting all results that follow.

1. **Discretionary vs. non-discretionary searches cannot be separated** - Some searches are driven largely by officer discretion, while others may be required by policy or circumstance (for example, searches tied to an arrest process or other mandatory procedures). The current data do not allow us to reliably tag each search as discretionary or non-discretionary. This matters because questions about racial or gender bias are most directly evaluated in decisions where officers have meaningful choice. When discretionary and non-discretionary searches are combined, observed differences may not map cleanly to officer decision-making.

2. **Arrests do not have a clear "productivity" measure in these data** - For searches, the data contain a straightforward outcome that can be used to evaluate search yield (e.g., whether contraband was found). For arrests, the dataset does not include a similarly consistent indicator of outcome quality or evidentiary strength (such as final case disposition, charge validity, prosecution outcomes, or convictions). Because of this, we cannot apply standard "outcome tests", like hit-rate or threshold-style tests, to arrests. We treat arrest rates in any analysis as the outcomes of searches themselves. Any arrest-based comparisons in this report should therefore be interpreted as differences in observed arrest rates - not as direct measures of arrest justification or efficiency.

3. **Limited information on pre-search justification and in-stop dynamics** - The dataset includes reasons why stops occurred, but it does not contain specific, structured data explaining why a search occurred (e.g., the documented basis for probable cause or reasonable suspicion for the search decision). It also lacks details about what happened during the stop - such as subject behavior, officer observations, or escalation factors - that may influence whether a search or arrest occurs. As a result,

even when we control for available factors (like location, time, stop reason, and demographics), unobserved circumstances may still explain part of the differences we observe across groups, and results should not be interpreted as definitive evidence of intent.

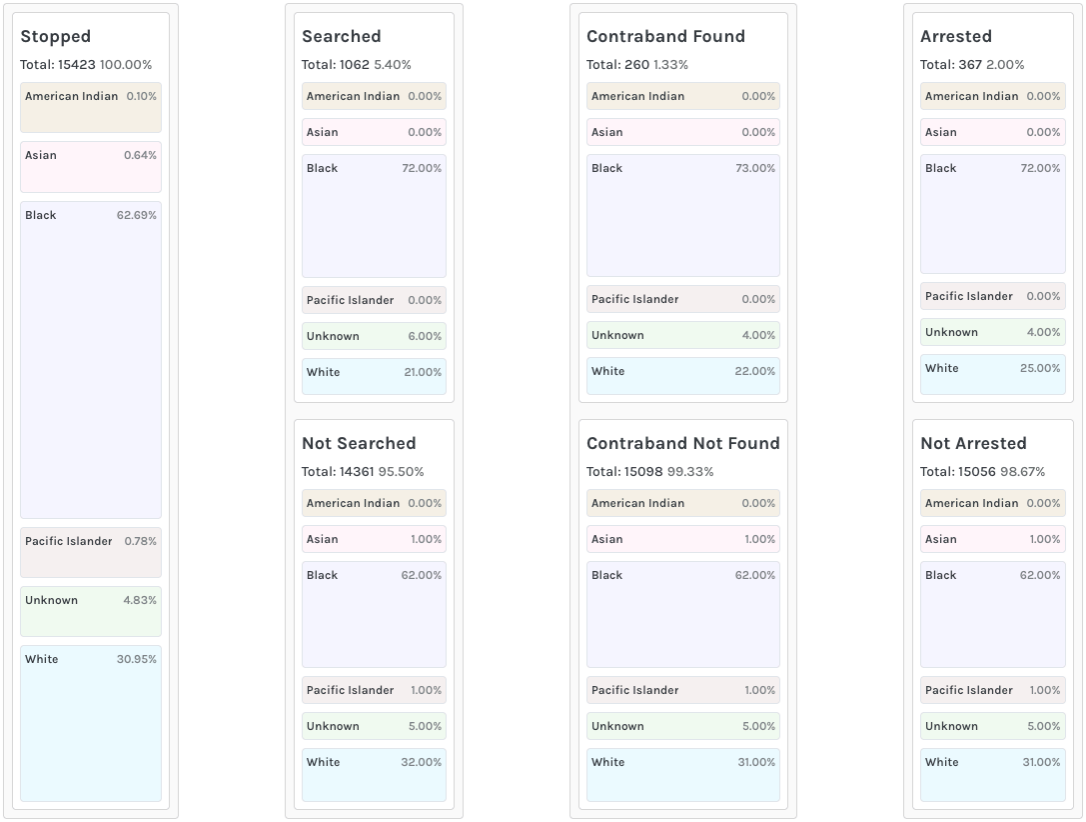
Stops Pipeline Analysis By Demographics

Introduction

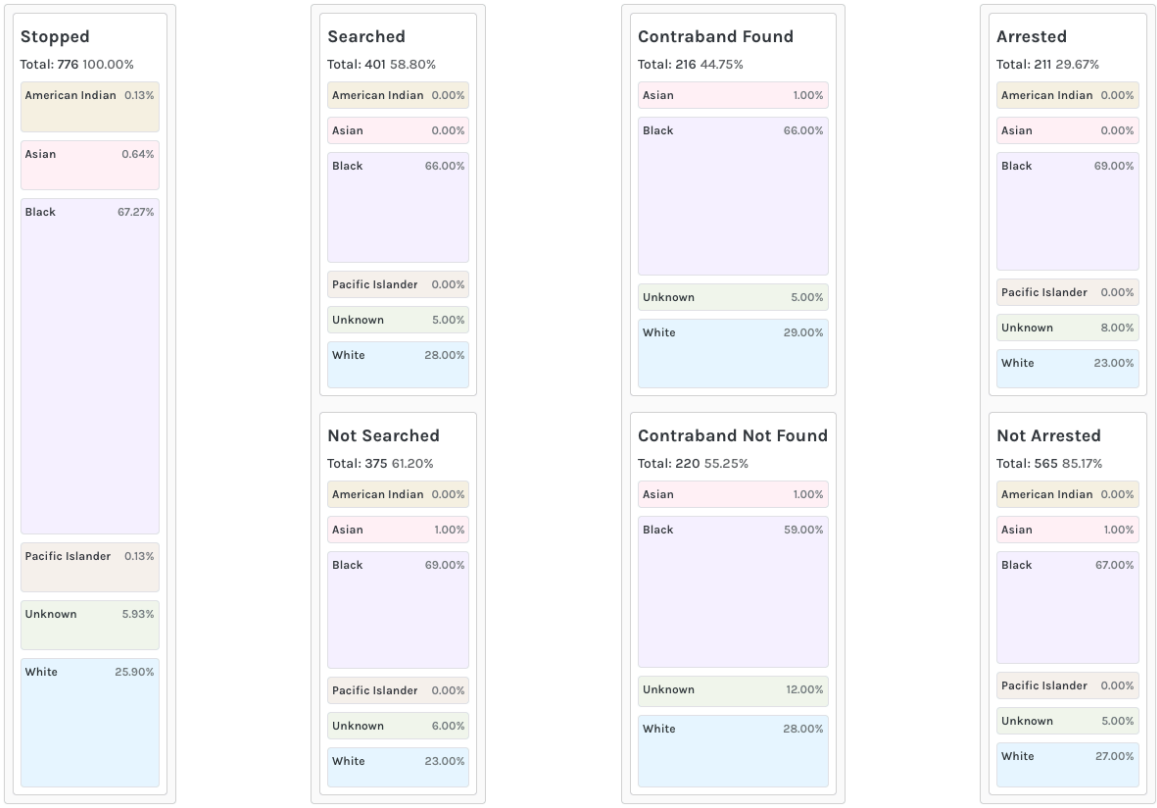
This section details the demographic composition of police interactions across the entire enforcement pipeline, tracking the proportional representation of subjects from initial stop through search, contraband recovery, and arrest. By segmenting data into **Traffic** and **Investigatory Stops**, the analysis isolates how racial and gender distributions shift at each decision point, establishing a baseline for understanding how enforcement intensity escalates or de-escalates for specific groups relative to their presence in the initial stop population.

Race

In 2024, **Traffic Stops** comprised 62.69% Black and 30.95% White subjects, along with 4.83% Unknown, 0.78% Pacific Islander, 0.64% Asian, and 0.10% American Indian subjects. During searches, the proportion of Black subjects increased to 72.00%, while the proportion of White subjects decreased to 21.00%. Contraband was subsequently found on 1.33% of the total stopped population - consisting of 73.00% Black subjects and 22.00% White subjects. With respect to the arrested population, arrests consisted of 72.00% Black subjects and 25.00% White subjects. The vast majority of traffic stops conclude without a search. Across such stops, the proportion for White subjects is 32% (compared to their search proportion of 21% and arrest proportion of 25%), while the proportion for Black subjects is 62% (compared to their search and arrest proportions of 72%).



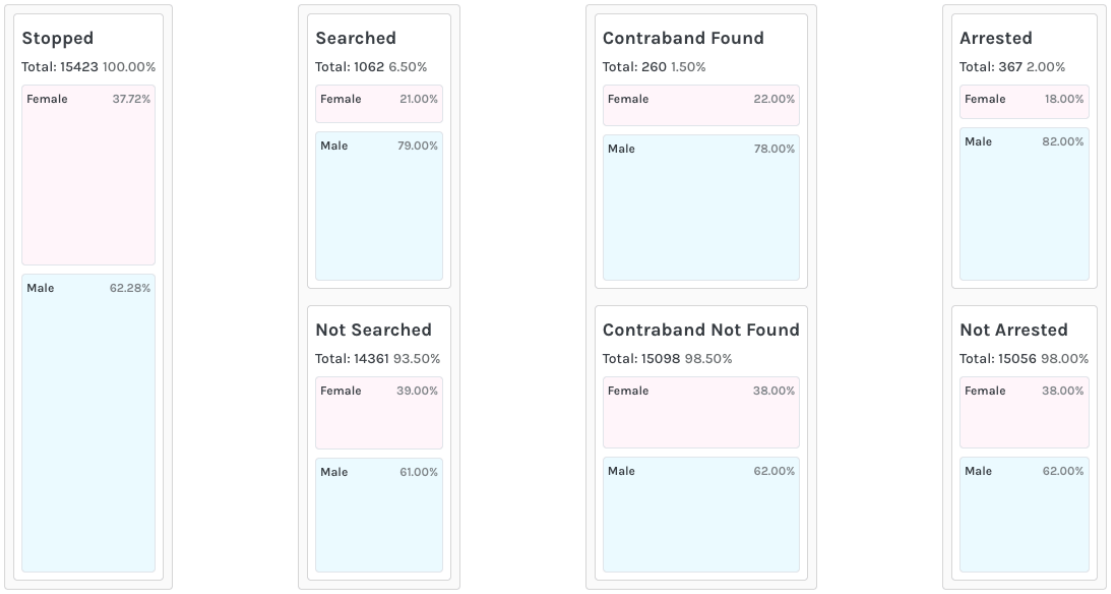
For **Investigatory Stops**, the demographic breakdown consists of 67.27% Black and 25.90% White subjects, along with 5.93% Unknown, 0.64% Asian, and 0.13% each for Pacific Islander and American Indian subjects. The demographic split at the search stage remained consistent with the initial stop population, recorded at 66.00% Black and 28.00% White, indicating a relatively stable Black-to-White ratio. These proportions remain consistent across investigatory stops that conclude without a search. When a subject was searched, contraband was found in a population consisting of 66.00% Black and 28.00% White subjects. With respect to the arrested population, arrests consisted of 69.00% Black subjects and 23.00% White subjects.



Gender

In 2024, **Traffic Stops** comprised 37.72% Female and 62.28% Male subjects. The proportion of Male subjects rose to 79.00%, while Female subjects decreased to 21.00% across the Search population. While females make up over a third of stops, they account for roughly one-fifth of searches. The outcomes of these interventions show that contraband was recovered on 1.5% of the stopped population - consisting of 22.00% Female and 78.00% Male subjects.

Arrests comprised 18.00% Female and 82.00% Male subjects. Male subjects were arrested at a proportion more than four times that of Female subjects. The vast majority of stops concluded without a search (93.50%) or an arrest (98.00%). For these non-arrest interactions, the demographic distribution remained closer to the initial stop population (38.00% Female and 62.00% Male for non-arrests).



For **Investigatory Stops**, the demographic breakdown consists of 19.72% Female and 80.28% Male subjects. This proportion remained relatively consistent at the search stage, recorded at 16.00% Female and 84.00% Male. This indicates a relatively stable male-to-female ratio across the initial stop and search stages. These searches yielded contraband on a population consisting of 18.00% Female and 82.00% Male subjects, and 19.00% Female and 81.00% Male subject arrests. Fifty-two percent of investigatory stops concluded without a search while 73% concluded without an arrest, with the non-arrested population mirroring the initial stop demographics at 20.00% Female and 80.00% Male.

Disparities and Biases By Demographics

Introduction

Overview

This section presents a quantitative analysis of police searches, focusing on differences in treatment across racial (Black/White) and gender (Female/Male) demographics. We examine Raw Disparities, which capture the unadjusted differences experienced by subjects, Controlled Disparities, which adjust for contextual factors such as stop circumstances and subject characteristics, and Bias-Free Disparities, which estimate the disparity that would remain if officers were responding uniformly to comparable signals of risk across groups. Together, these measures allow the report to distinguish between disparities arising from contextual differences and those that may reflect procedural variation in how enforcement decisions are made.

The report evaluates these disparities across Traffic Stops and Investigatory Stops, recognizing that observed differences may vary depending on the stop category. While the main text of the report focuses on Black vs. White and Female vs. Male disparities and biases only, results for all races can be found in Appendix B.

Definitions

Raw Disparities

Raw disparities describe differences in search rates across racial groups without applying any statistical controls. In other words, we directly compare the share of stops that result in a search for each group, treating all stops as equivalent. However, these don't reflect differences in the contexts in which groups are stopped (e.g., time of day or location). As a result, raw disparities should be viewed as a starting point rather than evidence of bias on their own.

Controlled Disparities

Controlled disparities estimate racial differences in search rates after accounting for observable factors that may affect search likelihood. Specifically, we control for variables such as district, zone, age, stop reason, stop time, stop year, stop type, and subject role. By holding these characteristics constant, the controlled disparity isolates the difference associated with race among otherwise similar stops. This approach reduces the risk of confounding and provides a clearer view of whether race remains predictive of searches once situational and demographic context is taken into account.

Bias Free Disparities

Bias-free disparities estimate what differences in search rates would look like if search decisions were driven only by observable risk signals and other contextual factors included in the analysis. To do this, we apply a set of bias tests that assess whether groups face different hit rates or implied decision thresholds after accounting for circumstances. These tests translate controlled disparities into the disparity that would remain under a consistent evidentiary standard – i.e., if officers responded uniformly to comparable observable indicators of risk. Comparing controlled and bias-free disparities helps distinguish differences that can be explained by context from differences that may reflect uneven decision standards across groups.

Statistical Significance

Throughout this report, we present all results alongside notes on their statistical significance. Statistical significance is a way of quantifying how confident we can be that an observed difference is not simply the product of random noise in the sample – i.e., it indicates whether the data provide strong evidence that a difference exists in the underlying population, not just in this particular dataset. We treat an estimate as statistically significant when its p-value is below 0.05. When we report results across many subsamples (i.e., conduct multiple hypothesis tests), we apply a Bonferroni correction[1] to reduce the risk of false positives that naturally increases as the number of tests grows. In those settings, we report whether estimates remain statistically significant using Bonferroni-adjusted p-values.

Overview of Bias Tests

To measure whether bias is present, this report conducts the Hit-Rate and Threshold tests to assess whether officers apply consistent evidentiary standards across groups and whether disparities reflect true underlying risk or differing decision thresholds.

[1] Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American statistical association, 56(293)*, 52-64.

Definitions

- **Hit-rate Test[2]:** Examines whether contraband is found at different rates across racial groups among those who are searched. If officers apply the same standard to all groups, hit rates should be comparable. Systematic differences in hit rates can suggest that some groups face a lower or higher bar for being searched.
- **Threshold Test[3]:** Estimates the implicit probability threshold at which officers decide to search individuals from different racial groups. If thresholds differ by race, it indicates that decision standards are not being applied uniformly, even if raw or controlled search rates appear similar.

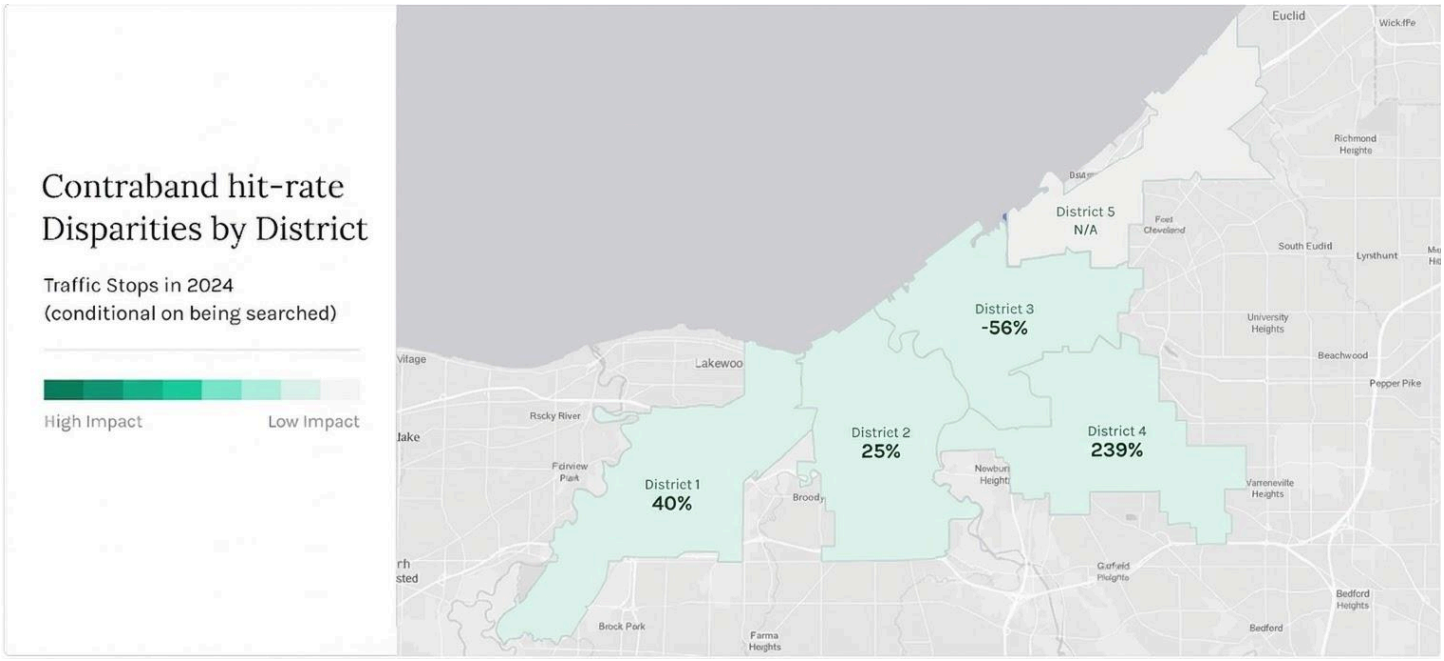
[2] Knowles, J., Persico, N., & Todd, P. (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of political economy*, *109*(1), 203-229.
[3] Simoiu, C., Corbett-Davies, S., & Goel, S. (2017). The problem of infra-marginality in outcome tests for discrimination. *The Annals of Applied Statistics, 11(3)*, 1193-1216.

Search Rate

Black, Traffic Stops

Raw, Controlled and Bias Free Disparities:

In the context of **Traffic Stops**, Black subjects are searched 64.6% more frequently, which reduces to 50.5% when accounting for factors such as district and zone of encounter, time and year of encounter, reason behind the stop, role of subject and age of subject. The controlled disparity is statistically significant. Searches during traffic stops do not fail the contraband hit-rate test or the threshold test against Black subjects, meaning there is no evidence of bias in Search decisions.



Deep Dive into Post Search Outcomes:

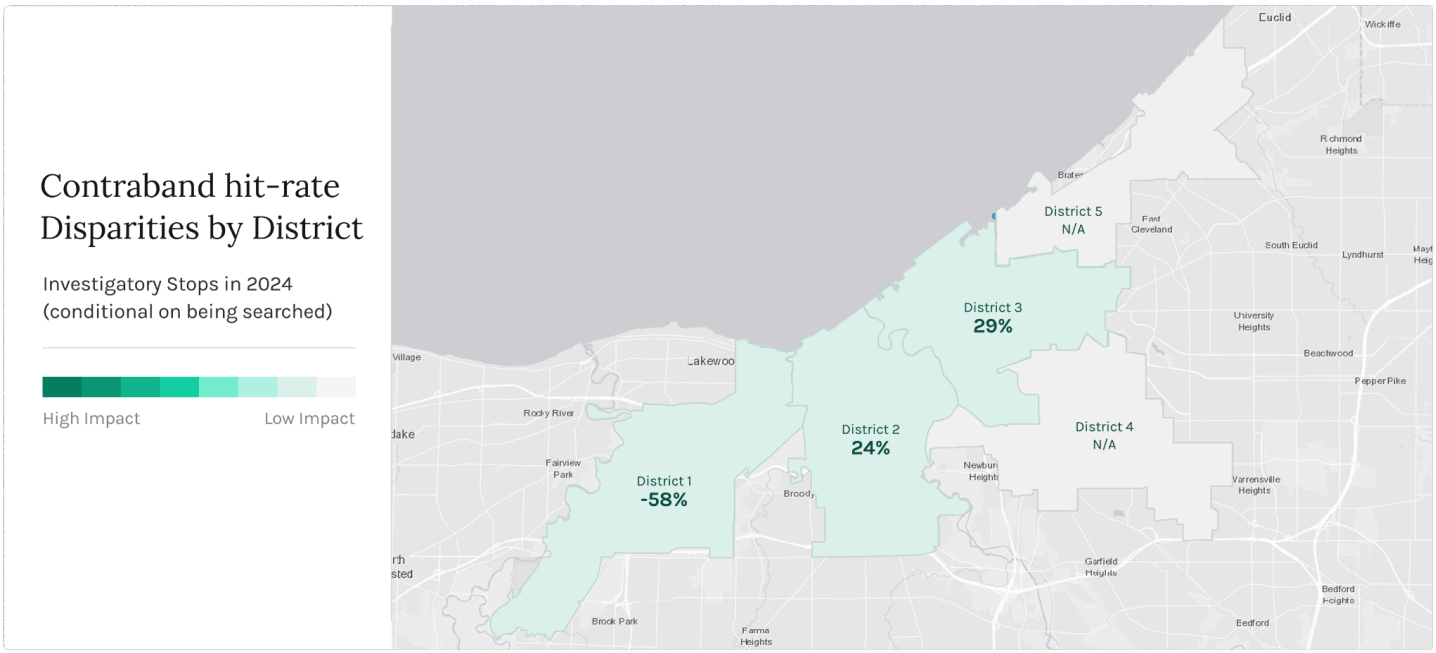
- Contraband Hit-Rate: The contraband hit-rate test shows that Black subjects have statistically similar contraband hit-rates as White subjects. Black subjects are 11.7% more likely to have contraband found during a search within similar contextual factors, but this difference is statistically insignificant at the 95% confidence level.
- Arrest Rate Conditional on Search: Conditional on being searched, Black subjects have statistically similar arrest rates as White subjects. Black subjects are arrested 12.5% less frequently than White subjects following searches with similar contextual factors, but this difference becomes statistically insignificant at a 95% confidence level.

Black, Investigatory Stops

Raw, Controlled and Bias Free Disparities:

In the case of **Investigatory Stops**, Black subjects are searched 10.3% less frequently than White subjects. Controlling for situational factors reveals that Black subjects are actually searched 13.5% more frequently in

similar stop contexts. As with traffic stops, searches during investigatory stops do not fail the contraband hit-rate test or the threshold test against Black subjects, meaning there is no evidence of bias in Search decisions.



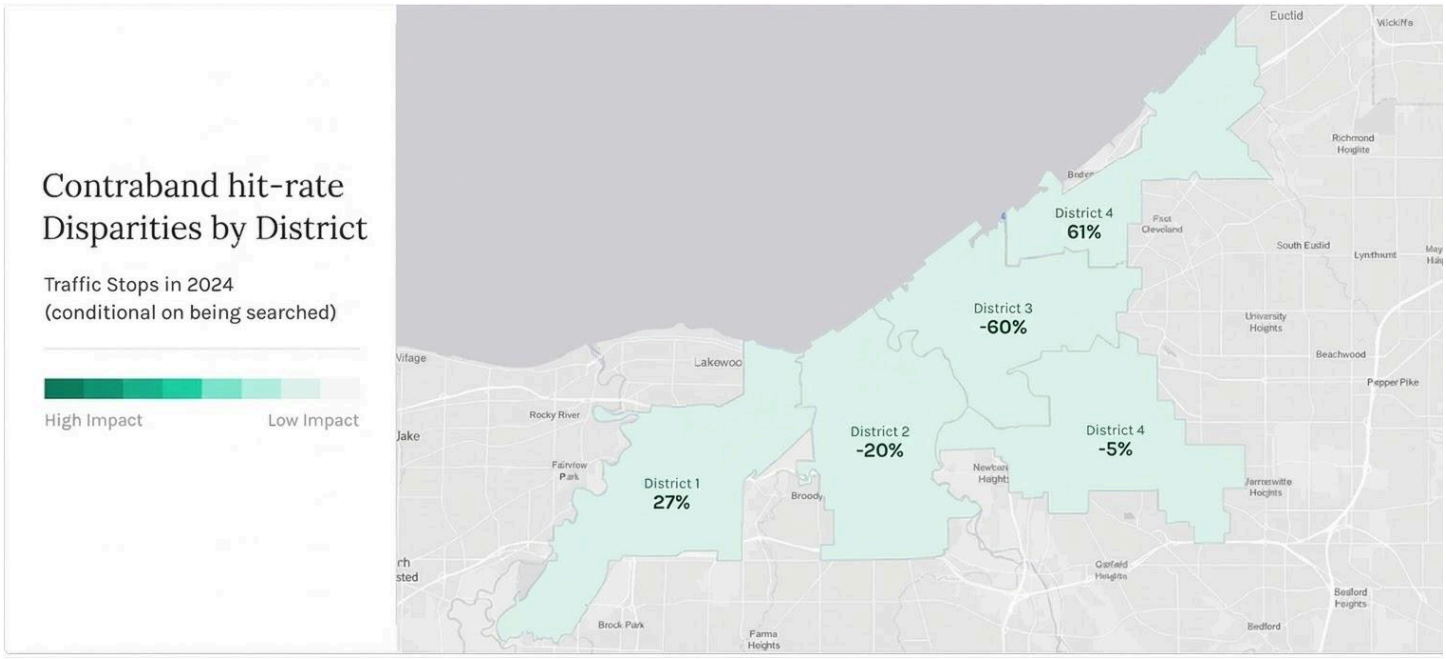
Deep Dive into Post Search Outcomes:

- Contraband Hit-Rate: The contraband hit-rate test shows that Black subjects have statistically similar contraband hit-rates as White subjects. Contraband is found on Black subjects 1.5% more frequently, but this becomes statistically insignificant at the 95% level.
- Arrest Rate Conditional on Search: Conditional on being searched, Black subjects have statistically similar arrest rates as White subjects (25.5% more compared to White subjects but statistically insignificant at the 95% level)

Female, Traffic Stops

Raw, Controlled and Bias Free Disparities:

In the context of **Traffic Stops**, Female subjects are searched 57.1% less frequently than Male subjects. Controlling for situational factors like district and zone of encounter, reason behind the stop, role of subject and age of subject, Female subjects are searched 47.1% less compared to Male subjects. Searches during traffic stops do not fail the contraband hit-rate test or the threshold test against Male subjects, meaning there is no evidence of bias in Search decisions



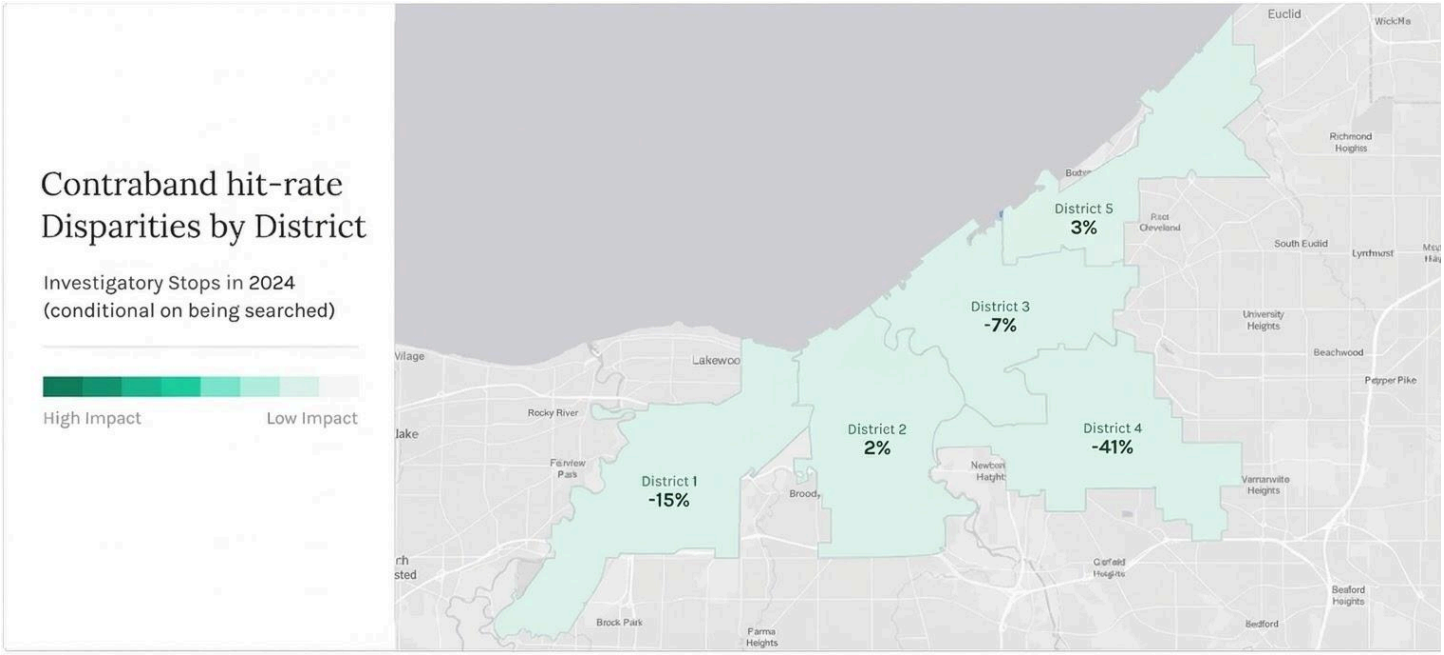
Deep Dive into Post Search Outcomes:

- Contraband Hit-Rate: The contraband hit-rate test shows that Female subjects have statistically similar contraband hit-rates as Male subjects. They are 8.7% more likely to have contraband found on them than similar Male subjects, but this becomes statistically insignificant at the 95% level.
- Arrest Rate Conditional on Search: Conditional on a search occurring, Female subjects are 29.3% less likely than Male subjects to be arrested after controlling for stop location, time of stop, stop reason, subject role, and age. This difference is statistically significant. Operationally, this means searches of Female subjects produce fewer arrests per search than searches of Male subjects. If arrests are the outcome being maximized through searches, the current search rate for female subjects appears inefficiently high relative to its arrest yield compared to males. Importantly, we are not distinguishing between discretionary and process-related searches in this analysis, which may be particularly relevant when looking at arrests conditional on search.

Female, Investigatory Stops

Raw, Controlled and Bias Free Disparities:

In the context of **Investigatory Stops**, Female subjects are searched 21.2% less frequently than Male subjects. Controlling for situational factors like district and zone of encounter, time and year of encounter, reason behind the stop, role of subject and age of subject, Female subjects are searched 30% less compared to Male subjects. Searches during traffic stops do not fail the contraband hit-rate test or the threshold test against Male subjects, meaning there is no evidence of systemic bias in Search decisions.



Deep Dive into Post Search Outcomes:

- Contraband Hit-Rate: The contraband hit-rate test shows that Female subjects have statistically similar contraband hit-rates as Male subjects. They are 8.1% less likely to have contraband found on them than similar Male subjects, but this is statistically insignificant at the 95% level.
- Arrest Rate Conditional on Search: Conditional on being searched, Female subjects have statistically similar arrest rates as Male subjects. They are 4.6% more likely to be arrested than similar Male subjects, but this is statistically insignificant at the 95% level.

Subsample Analysis By Demographics

Introduction

In this section, we examine how racial and gender differences in contraband hit-rates and arrest rates conditional on a search vary across districts. The goal is to assess whether searches are equally effective across groups - or whether effectiveness differs systematically depending on where the stop occurs - so we can identify the specific contexts driving the overall patterns.

We run subsample analyses because citywide averages can mask meaningful variation: there may be a small number of districts where disparities are concentrated and statistically significant. Subsample results help pinpoint where the differences are concentrated, facilitating more targeted and actionable operational responses.

District

Black, Traffic Stops

Across districts, the largest estimated racial differences in contraband hit-rates appear in Districts 3 and 4. City-wide, Black and White subjects have contraband found during Traffic Stops at statistically similar rates. In District 3, Black subjects are estimated to be 56% less likely than White subjects to have contraband found conditional on a search. In District 4, Black subjects are estimated to be 239% more likely than White subjects to have contraband found. The largest estimated racial difference in arrest rates conditional on a search are in District 3, where Black subjects are 46.6% less likely to be arrested after a search compared to White subjects. Note that none of these estimates remain statistically significant after adjusting for multiple hypotheses testing using a Bonferroni correction[4].

Group	Contraband Hit-Rate	Arrest Rate
District 1	40.10%	-6.00%
District 2	24.60%	24.80%
District 3	-55.90%	-46.60%
District 4	238.60%	-25.20%
District 5	N/A[5]	N/A

Black, Investigatory Stops

Across districts, the largest estimated differences vary by outcome. For contraband hit-rates conditional on a search, the biggest estimated gaps appear in District 1 (−58.0%) and District 3 (+28.6%), with District 2 showing a smaller positive difference (+23.9%). For arrest rates conditional on a search, the largest estimated differences appear in District 4 (+69.3%) and District 5 (+56.4%), followed by District 1 (+42.2%) and District 2 (+32.2%). None of these estimates remain statistically significant after adjusting for multiple hypothesis testing using a Bonferroni correction.

Group	Contraband Hit-Rate	Arrest Rate
District 1	-58.00%	42.20%
District 2	23.90%	32.20%
District 3	28.60%	N/A
District 4	N/A	69.30%

[4] Because we test disparities across multiple districts (and other subsamples), we adjust for multiple hypothesis testing using a Bonferroni correction to control the family-wise error rate. Specifically, if we conduct m simultaneous tests, we evaluate significance using an adjusted threshold $\alpha^{*} = \alpha/m$

[5] Cells with “N/A” do not have enough data to calculate a controlled disparity

District 5	N/A	56.40%

Female, Traffic Stops

Across districts, the estimated gender differences vary by outcome. For contraband hit-rates conditional on a search, the largest estimated gaps appear in District 5 (+60.9%) and District 3 (−60.1%), with smaller differences in District 1 (+26.9%), District 2 (−19.6%), and District 4 (−5.4%). For arrest rates conditional on a search, the estimated differences are negative in every district, with the largest gaps in District 5 (−60.9%) and District 3 (−25.7%), followed by District 2 (−22.9%), District 4 (−21.9%), and District 1 (−15.7%). None of these estimates remain statistically significant after adjusting for multiple hypothesis testing using a Bonferroni correction.

Group	Contraband Hit-Rate	Arrest Rate
District 1	26.90%	-15.70%
District 2	-19.60%	-22.90%
District 3	-60.10%	-25.70%
District 4	-5.40%	-21.90%
District 5	60.90%	-60.90%

Female, Investigatory Stops

In the context of investigatory stops for contraband hit-rates conditional on a search, the largest estimated gap appears in District 4 (−40.5%), followed by District 1 (−15.1%) and District 3 (−6.9%); District 2 (+2.3%) and District 5 (+2.7%) are near zero. For arrest rates conditional on a search, the largest estimated differences appear in District 5 (−87.7%), District 4 (+39.5%), and District 2 (+37.6%), with smaller negative differences in District 3 (−16.8%) and District 1 (−4.0%). With the exception of arrest rate differences in District 5, all estimates remain statistically insignificant after adjusting for multiple hypothesis testing using a Bonferroni correction.

Group	Contraband Hit-Rate	Arrest Rate
District 1	-15.10%	-4.00%
District 2	2.30%	37.60%
District 3	-6.90%	-16.80%
District 4	-40.50%	39.50%
District 5	2.70%	-87.70%

Key Report Findings

The findings in this report outline key patterns in enforcement outcomes within Cleveland. Through data-driven analysis, the Cleveland Division of Police can strengthen public transparency, improve decision-making across operational domains, and support a policing framework that reflects consistent application of enforcement practices.

Going forward, maintaining a commitment to continuous evaluation will be essential. The issues highlighted here are addressable and can be improved through updated policy design, targeted training, and collaborative engagement with community stakeholders. With sustained effort, Cleveland can become a national leader in using empirical analysis to refine law-enforcement operations.

- **Stops Pipeline Analysis:** In Traffic Stops, Black subjects make up 62.7% of stops but 72.0% of searches (White subjects fall from 31.0% of stops to 21.0% of searches), indicating a shift toward more searches among Black subjects within traffic interactions. In the same traffic context, Female subjects are 37.7% of stops but 21.0% of searches (Male subjects rise to 79.0% of searches). For Investigatory Stops, the Black/White composition is comparatively stable from stop to search (Black 67.3% at stops vs 66.0% at searches; White 25.9% vs 28.0%). Importantly, demographic composition shifts alone do not establish bias - they can also reflect differences in where/when stops occur, stop types, call-for-service patterns, deployment, and other contextual factors - so they must be interpreted alongside controlled analyses and post-search outcomes.

- **Disparities & Biases:** For Traffic Stops, Black subjects are searched 64.6% more often than White subjects (controlled disparity: 50.4%), but searches do not fail the hit-rate or threshold tests, and post-search outcomes (contraband hit-rate and arrest rate conditional on search) are statistically similar. For Investigatory Stops, the raw disparity is −10.3%, but the controlled disparity reverses to +13.5%, again with no bias-test failures and statistically similar post-search outcomes. This means we do not observe process bias happening in search decisions amongst the stopped population, despite the persistence of controlled disparities. Future analysis that delineates between discretionary and non-discretionary search decisions, as well as more structured data detailing why the search occurred, subject behavior, officer observations, and escalation factors would provide further detail on both the drivers of controlled disparities, and the relationship between those drivers and desired public safety outcomes.

  For gender, Female subjects are searched substantially less than Male subjects in both Traffic Stops (controlled: −47.1%) and Investigatory Stops (controlled: −30.0%), with contraband hit-rates generally similar; however, in Traffic Stops, conditional on a search, Female subjects are 29.3% less likely than Male subjects to be arrested (statistically significant).

- **Subsample Analysis:** District-level estimates show large variation in search “yield” measures - for example, Black vs. White contraband hit-rate, conditional on search, ranges from −55.9% (District 3) to +238.6% (District 4) in Traffic Stops - yet none of the Black/White district estimates remain statistically significant after Bonferroni adjustment. For Female vs. Male in Investigatory Stops, the largest estimated arrest-rate (conditional on search) difference appears in District 5 (−87.7%); aside from this District 5 arrest-rate result, the remaining district-level estimates are statistically insignificant after Bonferroni adjustment. Similar to the Disparities & Bias analysis, the Subsample analysis would benefit from further review that teases out differences in outcomes across discretionary and non-discretionary searches in order to more fully understand the context behind these District-level variances.

Appendix A

Data Used and Processing Notes

Source data

The analysis uses Cleveland Department of Police Stops data from 2024. The core input table includes incident and citation identifiers, stop metadata (date/time, district/zone, stop type, stop reasons), subject demographics (role, race/ethnicity, sex, age), search/contraband flags (subject and vehicle), and disposition fields (arrest/citation/warnings/other outcomes).

Some data-quality issues are noted here:

- **Incident IDs are not unique** (a single incident may contain multiple citation IDs which have the exact same subject demographics across all citations).
- **Citation IDs are not unique** (multiple records can share the same citation ID, and can differ on rejection status, rejection reason, and even subject demographics).
- Some records may be **invalid** (e.g., an example record with `Age=0.0`).

Transformation and feature engineering summary

The pipeline first flattens/joins the source inputs, then generates standardized features through a set of sequential feature generators.

Key steps include:

- **Column mapping (renaming only)** for fields like `stop_type` and `subject_role`.
- **Incident-level transformations** to produce analysis-ready fields such as `searched` and `search_type`, location fields (`district`, `zone`, `department`, `location`), binned fields (`age_range`, `stop_time_range`), and derived constructs like `veil_of_darkness` and consolidated `stop_reason`.
- **Subject demographic cleaning/mapping**, including age validation and standardized labels for gender, race, and ethnicity.
- **Contraband categorization** for both subject and vehicle contraband, using pattern matching on description text from the `contraband_2024` dataset and joining via citation and demographic keys.
- **Citation-level aggregation** flags (`was_citation_rejected`, `was_citation_accepted`) computed within each citation ID (`BrazosFormPK`).

Deduplication approach

Because duplicates can arise both within citations and within incidents (including suspicious repetitions of identical subject-demographic patterns), the pipeline applies a two-step “middle ground” deduplication:

1. **Citation-level deduplication**: keep one record per `BrazosFormPK` (keeping the last record).

2. **Pattern-based deduplication**: detect subject-identity patterns (incident + role + race + ethnicity + gender + age) occurring more than a tolerance threshold (default **> 5**) for a single incident, then deduplicate flagged patterns by keeping the record with the latest `entry_date`.

Appendix B

Controlled Disparities in Searches for All Races

Traffic Stops

Group	Search Rate	Sample size
Black	50.5%	9,669
Unknown	38.0%	745
Pacific Islander	-28.9%	121
American Indian	-256.9%	15
Asian	-44.8%	99

Investigatory Stops

Group	Search Rate	Sample size
Black	13.5%	522
Unknown	-23.8%	46
Pacific Islander	N/A	–
American Indian	N/A	–
Asian	1.3%	5

Controlled Disparities in Contraband Hit-Rates and Arrest Rates (conditional on a search) for All Races

Traffic Stops

Group	Contraband Hit-Rate	Arrest Rate	Sample size
Black	11.7%	-12.5%	760

Unknown	1.0%	-44.1%	69
Pacific Islander	N/A	N/A	–
American Indian	N/A	N/A	–
Asian	-101.5%	-27.2%	3

Investigatory Stops

Group	Contraband Hit-Rate	Arrest Rate	Sample size
Black	1.5%	25.5%	263
Unknown	-1.8%	47.5%	22
Pacific Islander	N/A	N/A	–
American Indian	N/A	N/A	–
Asian	N/A	N/A	–

Arrest Rate

In this section we report disparities in arrest rates by race and gender for all stops, regardless of whether a search happened or not.

Black, Traffic Stops

Raw, Controlled and Bias Free Disparities:

For **traffic stops**, Black subjects face a raw arrest disparity of approximately 44.3% compared to White subjects, which persists at 33.8% even when controlling for similar environmental and situational factors.

Black, Investigatory Stops

Raw, Controlled and Bias Free Disparities:

For investigatory stops, Black subjects see only a 18.1% difference, however, when adjusting for similar characteristics reveals that Black subjects are arrested 36.8% more frequently than their White peers.

Female, Traffic Stops

Raw, Controlled and Bias Free Disparities:

For **traffic stops**, Female subjects show a raw disparity of 63.1% less than Male subjects, which persists at 55.1% even when controlling for similar environmental and situational factors.

Female, Investigatory Stops

Raw, Controlled and Bias Free Disparities:

For investigatory stops, Female subjects see only a -1.8% difference, however, when adjusting for similar characteristics reveals that Female subjects are arrested -10.4% less frequently than their Male peers.

Contraband Type

Although the dataset includes a contraband “type” field, it is too sparse to support rigorous type-specific analysis. Among searches that resulted in contraband being found, the contraband type is precisely identified for only 42.16% of cases. For transparency, we still report the counts by contraband type in this appendix so readers can see what is - and is not - captured in the underlying data. The details of the type of contraband are listed here:

Type of Contraband	**Number of Observations**
Firearm	104
Marijuana	59
Crack	73
Alcohol	68
Drug Paraphernalia	64
Other Drugs	44
Cocaine	19
Pills	10
Methamphetamine	11
Cell Phone	9
Ecstacy	3
Fentanyl	2
Heroin	5
PCP	2

License Plate	2
Ammunition	2
Credit Card	1
Criminal Tools	3
Oxycodone	1
Knife	4
Xanax	1
Currency	17
Stolen Property	1